

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): AOKI, Makoto  
Serial No.: Not yet assigned  
Filed: March 24, 2004  
Title: INFORMATION PROCESSING SYSTEM AND METHOD  
Group: Not yet assigned

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

March 24, 2004

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, the applicant(s) hereby claim(s) the right of priority based on Japanese Patent Application No.(s) 2003-387942, filed November 18, 2003.

A certified copy of said Japanese Application is attached.

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



---

Carl I. Brundidge  
Registration No. 29,621

CIB/alb  
Attachment  
(703) 312-6600

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日                      2 0 0 3 年 1 1 月 1 8 日  
Date of Application:

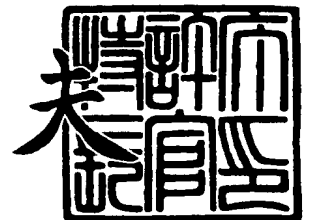
出 願 番 号                      特 願 2 0 0 3 - 3 8 7 9 4 2  
Application Number:  
[ST. 10/C] :                      [ J P 2 0 0 3 - 3 8 7 9 4 2 ]

出      願      人                      株式会社日立製作所  
Applicant(s):

2 0 0 4 年    3 月    5 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



出証番号    出証特 2 0 0 4 - 3 0 1 7 4 6 3

【書類名】 特許願  
【整理番号】 GM0309003  
【提出日】 平成15年11月18日  
【あて先】 特許庁長官殿  
【国際特許分類】 G06F 3/06  
【発明者】  
    【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所  
                        ソフトウェア事業部内  
    【氏名】 青木 誠  
【特許出願人】  
    【識別番号】 000005108  
    【氏名又は名称】 株式会社日立製作所  
【代理人】  
    【識別番号】 100075513  
    【弁理士】  
    【氏名又は名称】 後藤 政喜  
【選任した代理人】  
    【識別番号】 100084537  
    【弁理士】  
    【氏名又は名称】 松田 嘉夫  
【選任した代理人】  
    【識別番号】 100114236  
    【弁理士】  
    【氏名又は名称】 藤井 正弘  
【手数料の表示】  
    【予納台帳番号】 019839  
    【納付金額】 21,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1  
    【包括委任状番号】 0110326

**【書類名】 特許請求の範囲****【請求項 1】**

物理デバイスに論理的に設定された論理ユニットを備える記憶装置と、  
前記記憶装置に切り替えて接続され、前記記憶装置に対してデータ入出力を要求する複数の情報処理装置と、を備え、  
前記論理ユニットに対する通信経路となる複数のパスを経由してデータ入出力を要求する情報処理システムにおいて、  
前記情報処理装置は、  
データ入出力要求の結果からパスに発生した障害を検出する障害検出部と、  
所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも、前記記憶装置に接続される情報処理装置を切り替えるかを判定する切替評価部と、  
前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求する情報処理装置を切り替える切替部と、を備えることを特徴とする情報処理システム。

**【請求項 2】**

前記障害検出部は、前記検出した障害の種類を特定し、  
前記サーバ切替評価部は、前記特定された障害の種類によって、前記情報処理装置の切り替えの実行の判定基準となるパス数を変更することを特徴とする請求項 1 に記載の情報処理システム。

**【請求項 3】**

前記障害検出部は、パスの瞬断によって生じる瞬断エラーを特定し、  
前記サーバ切替評価部は、前記瞬断エラーによる情報処理装置の切り替えの実行の判定基準となるパス数を、他のエラーによる情報処理装置の切り替えの実行の判定基準となるパス数より大きく設定することを特徴とする請求項 2 に記載の情報処理システム。

**【請求項 4】**

前記障害検出部は、パスの性能の劣化によって生じる性能エラーを特定し、  
前記サーバ切替評価部は、前記性能エラーによる情報処理装置の切り替えの実行の判定基準となるパス数を、他のエラーによるフェイルオーバーの実行の判定基準となるパス数より小さく設定することを特徴とする請求項 2 に記載の情報処理システム。

**【請求項 5】**

前記障害検出部は、前記検出した障害の種類を特定し、  
前記サーバ切替評価部は、前記特定された複数種類の障害の組み合わせに基づいて前記情報処理装置の切り替えの実行を判定することを特徴とする請求項 1 に記載の情報処理システム。

**【請求項 6】**

前記サーバ切替評価部は、パスに発生した種類の障害に応じて定められた点数の集計結果に基づいて、前記情報処理装置の切り替えを実行するかを判定することを特徴とする請求項 5 に記載の情報処理システム。

**【請求項 7】**

前記サーバ切替評価部は、データ入出力要求の負荷状態に応じて前記情報処理装置の切り替えの実行の判定基準となるパス数を変更することを特徴とする請求項 1 に記載の情報処理システム。

**【請求項 8】**

前記サーバ切替評価部は、データ入出力要求の負荷状態が高くなる時間帯には、前記情報処理装置の切り替えの実行の判定基準となるパス数を、他の時間帯より小さく設定することを特徴とする請求項 1 に記載の情報処理システム。

**【請求項 9】**

物理デバイスに論理的に設定された論理ユニットを備える記憶装置と、  
前記記憶装置に切り替えて接続され、前記記憶装置に対してデータ入出力を要求する複数の情報処理装置と、を備え、  
前記論理ユニットに対する通信経路となるパスを経由してデータ入出力を要求する情報

処理システムにおいて、

前記情報処理装置は、

前記記憶装置に送信されるデータ入出力要求を割り当てるパスを選択するパス選択部と

、  
前記パス選択部によって選択されたパスに対して発生したデータ入出力要求を送出する  
I/O送受信部と、

正常に終了したデータ入出力要求の処理状態を集計する稼動統計管理部と、

前記パスに発生した障害を検出し、該障害の種類を特定し、パス毎、エラーの種類毎に  
エラーの検出回数を集計する障害管理部と、

所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも  
、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の切り替えを実行す  
るかを判定する切替評価部と、

前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求  
する情報処理装置を切り替える切替部と、を備えることを特徴とする情報処理システム。

【請求項 10】

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記記憶  
装置に対して切り替えて接続され、前記論理ユニットに対する通信経路となる論理パスを  
経由してデータ入出力を要求する情報処理装置において、

データ入出力要求の結果からパスに発生した障害を検出する障害検出部と、

所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも  
、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の切り替えを実行す  
るかを判定する切替評価部と、

前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求  
する情報処理装置を切り替える切替部と、を備えることを特徴とする情報処理装置。

【請求項 11】

前記障害検出部は、前記検出した障害の種類を特定し、

前記サーバ切替評価部は、前記特定された障害の種類によって、前記情報処理装置の切  
り替えの実行を判定基準となるパス数を変更することを特徴とする請求項 10 に記載の情  
報処理装置。

【請求項 12】

前記障害検出部は、前記検出した障害の種類を特定し、

前記サーバ切替評価部は、前記特定された複数種類の障害の組み合わせに基づいて前記  
情報処理装置の切り替えの実行を判定することを特徴とする請求項 10 に記載の情報処理  
装置。

【請求項 13】

前記サーバ切替評価部は、データ入出力要求の負荷状態に応じて前記情報処理装置の切  
り替えの実行の判定基準となるパス数を変更することを特徴とする請求項 10 に記載の情  
報処理装置。

【請求項 14】

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記論理  
ユニットに対する通信経路となる論理パスを經由してデータ入出力を要求する情報処理装  
置において、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の制御方  
法であって、

データ入出力要求の結果からパスに発生した障害を検出し、

所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも  
、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の切り替えを実行す  
るかを判定し、

前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求  
する情報処理装置を切り替えることを特徴とする制御方法。

【請求項 15】

前記検出した障害の種類を特定し、  
前記特定された障害の種類によって、前記情報処理装置の切り替えの実行の判定基準となるパス数を変更することを特徴とする請求項 14 に記載の制御方法。

【請求項 16】

前記検出した障害の種類を特定し、  
前記特定された複数種類の障害の組み合わせに基づいて前記情報処理装置の切り替えの実行を判定することを特徴とする請求項 14 に記載の制御方法。

【請求項 17】

データ入出力要求の負荷状態に応じて前記情報処理装置の切り替えの実行の判定基準となるパス数を変更することを特徴とする請求項 14 に記載の制御方法。

【請求項 18】

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理装置において、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の制御方法であって、

データ入出力要求の結果からパスに発生した障害を検出して、障害管理テーブルに記憶し、

前記障害管理テーブルを参照して、パス毎に障害検出数が所定の判定値を超えているか否かを判定し、

前記障害検出数と前記所定の判定値との比較結果に基づいて、エラーが発生したパスを特定し、

前記エラーが発生していると特定されたパスの数と、所定の閾値とを比較した結果に基づいて、前記情報処理装置を切り替える指示を発行することを特徴とする制御方法。

【請求項 19】

物理デバイスに論理的に設定された論理ユニットを備えた記憶装置に対して、前記論理ユニットに対する通信経路となる論理パスを経由してデータ入出力を要求する情報処理装置を機能させるプログラムであって、

データ入出力要求の結果からパスに発生した障害を検出する手段と、

所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の切り替えを実行するかを判定する手段と、

前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求する情報処理装置を切り替える手段として機能させることを特徴とするプログラム。

【請求項 20】

前記検出した障害の種類を特定する手段と、

前記特定された障害の種類によって、前記情報処理装置の切り替えの実行の判定基準となるパス数を変更する手段として機能させることを特徴とする請求項 19 に記載のプログラム。

【請求項 21】

前記検出した障害の種類を特定する手段と、

前記特定された複数種類の障害の組み合わせに基づいて前記情報処理装置の切り替えの実行を判定する手段として機能させることを特徴とする請求項 19 に記載のプログラム。

【請求項 22】

データ入出力要求の負荷状態に応じて前記情報処理装置の切り替えの実行の判定基準となるパス数を変更する手段として機能することを特徴とする請求項 19 に記載のプログラム。

**【書類名】 明細書**

**【発明の名称】** 情報処理システム、情報処理装置、情報処理装置の制御方法及びプログラム

**【技術分野】****【0001】**

本発明は、情報処理システム、情報処理装置、情報処理装置の制御方法、及びプログラムに関し、特に、記憶装置の障害の検出に関する。

**【背景技術】****【0002】**

近年、情報処理システムで取り扱われるデータ量が増大している。大容量化した記憶装置では記憶容量に見合うだけの入出力性能と信頼性を確保する必要がある。そのため記憶装置への論理パス（入出力経路）を多重化し、記憶装置へのデータ入出力要求（I/O要求）を論理パスに適宜割り当ててものが開発されている。

**【0003】**

この論理パスに障害が発生したときに、全ノードの論理ディスク管理テーブルを書き換えることによってI/Oパスの切り替えを行うことによって、I/Oパス切り替えのために要する時間を短縮する技術が提案されている（例えば、特許文献1参照。）。

**【0004】**

また、システムを構成し、切り離し可能な各部位毎に障害検出回数を記憶するカウンタを設け、またそのカウンタ値を総合的に判定する論理を具備することによって、故障部位を高精度に特定し、故障部位をシステムより排除し、正常部位を用いたシステム動作を継続可能とする技術も提案されている（例えば、特許文献2参照。）。

**【特許文献1】** 特開2002-49575号公報

**【特許文献2】** 特開平11-296311号公報

**【発明の開示】****【発明が解決しようとする課題】****【0005】**

前述した従来の技術では、複数設けられた全てのパスに障害が発生して、何回かリトライを行ってI/O処理が行えないことが検出されてからサーバを切り替えるものであり、障害検出期間中のスループットの低下については考慮されていなかった。

**【0006】**

本発明は、全パスに障害が検出される前に予防的にサーバを切り替えることによって、検出期間中のスループットの低下を防止することを目的とする。

**【課題を解決するための手段】****【0007】**

本発明は、物理デバイスに論理的に設定された論理ユニットを備える記憶装置と、前記記憶装置に切り替えて接続され、前記記憶装置に対してデータ入出力を要求する複数の情報処理装置と、を備え、前記論理ユニットに対する通信経路となる複数のパスを経由してデータ入出力を要求する情報処理システムにおいて、前記情報処理装置は、データ入出力要求の結果からパスに発生した障害を検出する障害検出部と、所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の切り替えを実行するかを判定する切替評価部と、前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求する情報処理装置を切り替える切替部と、を備えることを特徴とする。

**【発明の効果】****【0008】**

本発明によると、サーバを切り替える際のパス障害検出期間中のスループットを向上させることができる。

**【発明を実施するための最良の形態】****【0009】**

以下、本発明の実施の形態を図面を参照して説明する。

【0010】

図1は、本実施の第1の形態の情報処理システムの全体構成を示すブロック図である。

【0011】

情報処理システムは、複数のサーバ（情報処理装置）100a、100bがクラスタ構成をとっているクラスタシステム10、及び、少なくとも1台の記憶装置200を含んで構成される。なお、本実施の形態においては、1台のクラスタシステム10と1台の記憶装置200とを備える情報処理システムについて説明するが、複数台のクラスタシステム10及び複数台の記憶装置200を備える情報処理システムであってもよい。

【0012】

サーバ100は、CPU（Central Processing Unit）やメモリ等を備え、クライアント端末20からの指示によって、記憶装置200に記憶されたデータにアクセスしながら、各種アプリケーションプログラムを実行する。また、サーバ100には、クラスタシステム10内での当該サーバの動作を管理し、サーバを切り替えるフェイルオーバー処理を行うクラスタ管理部115（クラスタソフトウェア）115が設けられている。また、サーバ100は、複数のパスを論理的に管理し、アプリケーションプログラム110からのI/O要求を発行するパスを定めるパス管理部120を備えている。

【0013】

サーバ100には、ネットワーク（例えば、SAN（Storage Area Network））を介して記憶装置200が接続されている。サーバ100と記憶装置200との間のSANを介する通信は、一般にファイバチャネルプロトコル（Fibre Channel Protocol）に従って行われる。すなわち、サーバ100からは、記憶装置200に対して、ファイバチャネルプロトコルに従ってブロック単位のデータアクセス要求が送信される。なお、サーバ100と記憶装置200とは、必ずしもSANで接続される必要はなく、LAN等のネットワークを介して接続されてもよいし、SCSI（Small Computer System Interface）インターフェースによって直接接続されてもよい。

【0014】

サーバ100と記憶装置200との間を接続するSANには、サーバ100と記憶装置200とを接続するハードウェアによって物理的に構成される通信経路である物理パス300が設定されている。サーバ100は、複数の物理パス300を介して記憶装置200に記憶されているデータにアクセスする。パス管理部120は、記憶装置200へのI/O要求を送信する物理パス300を多重化して、帯域幅を拡大することによって、データアクセスにおけるボトルネックを解消し、データ入出力処理性能を向上させる。なお、パス管理部120は、物理パス300に対応して設定される論理パスによって、物理パス300を管理する。

【0015】

また、サーバ100は、ネットワーク（例えば、LAN：Local Area Network）を介して、記憶装置クライアント端末20と接続されている。

【0016】

記憶装置200は、複数の論理ユニット（LU：Logical Unit）220を備え、サーバ100から送信されてくるI/O要求に応じてデータの入出力を処理する。論理ユニット220は、記憶装置200に備わる記憶資源（物理デバイス）を論理的に設定した記憶領域である。図1において、記憶装置200は、二つの論理ユニットによって構成されているが、3以上の論理ユニットを備えてもよい。記憶資源としては、ハードディスク装置（ディスクアレイ装置）の他、フレキシブルディスク装置や半導体記憶装置等、様々な記憶媒体を用いることができる。

【0017】

なお、本実施の形態では、サーバ100がコンピュータ（パーソナルコンピュータ、ワークステーション、メインフレームコンピュータ等）である場合について説明したが、サーバ100が、ルータ、ハブ等のネットワーク装置であってもよい。



**【0018】**

図2は、本発明の第1の実施の形態のサーバ100の詳細を表した機能ブロック図である。

**【0019】**

サーバ100は、クラスタ管理部115、サーバ切り替え部123、パス管理ドライバ130備えている。また、サーバ100には、ホスト・バス・アダプタ（HBA）及びHBAドライバで構成されるインターフェース（IF）140が設けられている。

**【0020】**

パス管理ドライバ130は、IO送受信部131、パス選択部132、障害検出部133、障害管理部134、稼動統計管理部135及びサーバ切替評価部136が設けられている。

**【0021】**

IO送受信部131は、パス管理部120が受け取ったIO要求を、パス選択部132によって選択されたパス（インターフェース140）に送出し、インターフェース140が受け取ったIO終了をアプリケーション110に通知する。

**【0022】**

パス選択部132は、サーバ100と記憶装置200との間の物理パス300を論理的に管理する。また、パス選択部132は、アプリケーション110が発行したIO要求を割り当てるIF140を決定して、当該IO要求に使用されるパス300を選択する。そして、IO要求を各物理パス300へ適宜振り分けることによって負荷分散処理も行う。さらに、パス選択部132は、ある物理パス300に障害が発生した場合に当該物理パス300を切り離して、正常な物理パス300を通じて記憶装置200へのデータアクセスを継続する機能も有している。これにより、物理パスの障害によるサービス停止を回避しシステムの信頼性を高めている。

**【0023】**

障害検出部133は、IO要求に対して記憶装置200が発行したIO終了通知によって、IO処理が正常に終了したかを監視する。障害検出部133は、IO処理が異常終了したパスID、パスの状態（オンラインか、オフラインか）、障害の発生原因、異常終了が通知された時刻を記憶するIO管理テーブルを保持している。障害検出部133が、IO処理が異常に終了したことを検出すると、IO管理テーブルに障害が発生したパスに関する情報を記憶して、IO管理テーブルを更新する。

**【0024】**

また、障害検出部133は、アプリケーション110からのIO要求及び記憶装置200からのIO終了通知から、正常に終了したIO要求のデータ量、当該IO要求に要した処理時間を取得する。

**【0025】**

障害管理部134は、パス障害管理テーブル（図7）を保持している。障害管理部134は、IO管理テーブルを読み出すことによってパスの障害情報を取得する。そして、取得したパスの障害情報をパス障害管理テーブルに記憶し、パス毎、エラーの種類毎に計数されるエラーの検出回数（検出値）を更新する。

**【0026】**

稼動統計管理部135は、稼動統計管理テーブル（図9）を保持している。稼動統計管理部135は、障害検出部133が取得したIO要求及びIO終了通知から、正常に終了したIO要求のデータ量、当該IO要求に要した処理時間を取得し、稼動統計管理テーブルの該当パス及び該当期間のデータ量（バイト数）及び時間のデータを更新する。

**【0027】**

サーバ切替評価部136は、パスに発生したエラーの検出結果に基づいて、クラスタ内のサーバを切り替えるフェイルオーバを実行するかを決定し、その結果をサーバ切替指示部116に送る。このフェイルオーバを実行するかを判定するため、サーバ切替評価部136は、パスに発生するエラーを解析して、以下の3種類のエラー（瞬断エラー、累計エ

ラー、性能エラー)を解析する。

#### 【0028】

断線エラーは、同一パスで一つの I/O においてエラーが発生したが、続けて発行される I/O は正常に終了した場合に、瞬断エラーと判定される。瞬断エラーは、光ファイバの摩擦や、電源の不安定等の要因によって、伝送経路が不安定となった場合に生じる。

#### 【0029】

累計エラーは、一定期間内(例えば、一月間)に、所定回数を超えるエラーが発生した場合、累計エラーと判定される。なお、一定期間のエラー回数によって累計エラーを判定するのではなく、エラーを計数する期間を定めずに、一定回数のエラーが発生したら累計エラーと判断してもよい。また、連続した所定回数の I/O 要求の中でエラーとなった回数が所定の閾値を超えたり、全 I/O 要求のうち所定割合の I/O 要求がエラーとなったら累計エラーと判断してもよい。

#### 【0030】

性能エラーは、このパスが設置された状態と比較して、所定の値だけ性能が劣化した場合に、性能エラーと判定される。後述するように、稼動統計管理部 135 において I/O 処理のデータ量と処理時間を取得し、設置当初よりレスポンスが 60%劣化した場合に性能エラーと判定することができる。

#### 【0031】

IF140 は、記憶装置 200 との間で通信を行うためのインタフェースであり、ホスト・バス・アダプタ(HBA)及び HBA ドライバを含んでいる。HBA は、例えば、SCSI (Small Computer System Interface) アダプタ、ファイバチャネルアダプタ等の HBA (Host Bus Adapter) である。IF140 は、パス 300 に固有に設けられており、パス選択部 132 によって割り当てられた I/O 要求を処理し、記憶装置 200 に対してアクセスする。すなわち、サーバ 100 は、IF140 によって、記憶装置 200 との間でデータ入出力要求(I/O 要求)等を授受することができる。

#### 【0032】

図 3 は、本発明の第 1 の実施の形態の障害検出部の動作を示すブロック図である。

#### 【0033】

アプリケーション 110 は、記憶装置 200 に記憶されているデータにアクセスするために、パス管理部 120 に入出力を要求する。パス管理部 120 は、I/O 要求受付部 121 によって、アプリケーション 110 によって発行される I/O 要求を受け取る。受け取った I/O 要求は、パス管理ドライバ 130 に送られ、物理パス 301~308 が選択されて、I/O 要求が選択されたパスに対して送信される。

#### 【0034】

物理パスがオンライン状態(稼動時)であれば、アプリケーション 110 によって発行された I/O 要求は、記憶装置 200 によって処理され、I/O 処理が正常に終了した旨の I/O 終了通知が返送される。この I/O 処理通知は、I/O 終了通知部 122 によってアプリケーション 110 に通知される。

#### 【0035】

しかし、物理パスがオフライン状態(障害時)では、アプリケーション 110 によって発行された I/O 要求は、記憶装置 200 によって処理されることなく、記憶装置 200 までの経路においてタイムアウトし、パス管理ドライバ 130 に対して I/O 処理が異常終了した旨を I/O 終了通知が返送される。パス管理ドライバ 130 は、I/O 処理が異常終了した旨を検出すると、他のパスが選択されて、I/O 要求が記憶装置 200 に対して送信される。そして、I/O 要求が記憶装置 200 によって処理されると、I/O 処理が正常に終了した旨の I/O 終了通知が返送される。

#### 【0036】

例えば、図 3 に示すように、パス 302 に対して I/O 要求を送信したが、このときパス 302 に障害が生じていた場合、障害検出部 133 がこれを検出して、パス選択部 132 が他のパス 303 を選択して、I/O 要求を記憶装置 200 に対して送信する。そして、記

憶装置 200 から返送された I/O 終了通知は、アプリケーション 110 に対して送信される。

#### 【0037】

前述したパスのエラーは、主にパスを構成する伝送経路の一時的な断線によって生じる障害である。例えば、光ファイバの摩耗がによって瞬断が生じる。よって、同時期に設置された光ファイバは同時期に摩耗して一時的な断線状態（瞬断状態）となり、I/O 処理に一時的な障害が発生する。そして、何ら措置がとられないと、完全な断線状態に至って I/O 障害となる。

#### 【0038】

サーバ切替部 116 は、サーバ切替評価部 136 による決定に基づいて、サーバを切り替えるクラスタフェイルオーバーを実行させるためのサーバ切替指示をクラスタ管理部 115 に発行する。

#### 【0039】

図 4 は、本発明の第 1 の実施の形態のクラスタ切替判定に用いるテーブルの例を説明する図であり、8 本のパスが設けられている場合に、サーバ切替部 116 が、前述した 3 種類のエラーの判定結果に基づいて、サーバを切り替える判定基準を示す。

#### 【0040】

テーブルの上段に示すケース 1 では、全パスのうち半数以上のパスに瞬断エラーが発生した場合に（正常なパスが半数未満となった場合に）、サーバを切り替えることを示す。なお、理解の容易のため図 4 では、パス 1 ～パス 4 に瞬断エラーが発生した場合を示したが、エラーが発止したパスが連続しない他の 4 本のパスの組み合わせである場合にもフェイルオーバーを行うと判定される。

#### 【0041】

テーブルの中段に示すケース 2 では、全パスのうち 70% 以上のパスに累積エラーが発生した場合に（正常なパスが 30% 未満となった場合に）、サーバを切り替えることを示す。なお、理解の容易のため図 4 では、パス 1 ～パス 6 に累積エラーが発生した場合を示したが、エラーが発止したパスが連続しない他の 6 本のパスの組み合わせである場合にもフェイルオーバーを行うと判定される。

#### 【0042】

テーブルの下段に示すケース 3 では、全パスのうち 80% 以上のパスに性能エラーが発生した場合に（正常なパスが 20% 未満となった場合に）、サーバを切り替えることを示す。なお、理解の容易のため図 4 では、パス 1 ～パス 7 に性能エラーが発生した場合を示したが、エラーが発止したパスが連続しない他の 7 本のパスの組み合わせである場合にもフェイルオーバーを行うと判定される。

#### 【0043】

なお、一つのパスに複数のエラーが発生していると判定される場合には、スループットの低下への影響が大きく、情報処理システムが停止に至る可能性が高い重大なエラーを当該パスに生じたエラーであると判定する。図 4 に示す場合では、瞬断エラーは光ファイバの摩耗が主たる原因であり、同時期に設置された光ファイバは同時期に摩耗して故障する可能性が高いため、半数のパスにエラーが発生した場合にフェイルオーバーを実行して、早めにサーバを切り替えて、スループットの低下を防止している。一方、性能エラーの場合には、I/O 処理に時間がかかるだけで、I/O の処理が行えない訳ではないので、多くのパスに性能エラーが発生するまでフェイルオーバーを行うことなく、現用のサーバ及びパスにて稼働させるようにしている。

#### 【0044】

図 5 は、本発明の第 1 の実施の形態のサーバ切替処理のフローチャートである。

#### 【0045】

まず、アプリケーション 110 が発行し I/O 要求受付部 121 が受信した I/O 要求は、パス選択部 132 が選択したパス 301 ～308 に対して、I/O 送受信部 131 によって送信される。そして、記憶装置 200 による I/O 処理が完了すると、I/O 処理が正常に終

了した旨の I O 終了通知が、記憶装置 200 から発行され、I O 送受信部 131 及び I O 終了通知部 122 を経由して、アプリケーション 110 に返される (S501)。なお、I O 処理が正常に終了しなかった場合には、I O 処理が異常終了した旨の I O 終了通知は発行される。

#### 【0046】

I O 処理が完了すると、当該 I O 処理に関する I O 障害 (エラー) を検出する (S502)。そして、当該検出された障害に関する情報を障害管理テーブルへ記載して、障害管理テーブルを更新し (S503)、ステップ S505 に進む。一方、I O 障害「無」と判定されると、当該パスの I O 処理の結果を稼動統計テーブルへ記載し、稼動統計テーブルの当該パスの性能情報を更新し (S504)、ステップ S505 に進む。

#### 【0047】

ステップ S505 では、図 4 において前述した方法 (又は、図 12 において後述する方法) によって、全パスの障害状態を評価する。そして、所定の閾値との比較結果に基づいてサーバ切替の必要性を判定する (S506)。

#### 【0048】

図 6 は、本発明の第 1 の実施の形態の障害管理テーブル更新処理のフローチャートであり、サーバ切替処理 (図 5) のステップ S503 において実行される。

#### 【0049】

まず、障害検出部 133 が、記憶装置 200 等から発行された I O 終了通知によって、I O 処理が正常に行われたか、I O 処理の途中でエラーが生じたかを評価する (S601)。

#### 【0050】

そして、I O 障害の有無を判定する (S602)。その結果、I O 障害「無」と判定されると、累計エラーに関する処理 (S603～S606) を実行することなく、ステップ S607 へ移行する。一方、I O 障害「有」と判定されると、ステップ S603 へ移行して、累計エラーに関する処理 (S603～S606) を実行する。

#### 【0051】

累計エラーに関する処理では、まず、障害管理テーブル内の当該パスの累計エラー欄における開始時刻と監視期間を参照して、現在時が所定の監視時間範囲内であるかによって、所定の監視時間内に生じた I O 障害であるかを判定する (S603)。その結果、所定の監視時間内に生じた I O 障害であれば、ステップ S606 に移行し、累積エラーの検出値にエラー回数 (1 回) を加算し、この処理を終了する。

#### 【0052】

一方、ステップ S603 による判定の結果、所定の監視時間外に発生したエラーであると判定されると、現在時刻を監視の開始時刻に設定し (S604)、累計エラーの検出値を "0" にして (S605)、新たな監視時間を開始する。そして、累積エラーの検出値にエラー回数 (1 回) を加算し (S606)、この処理を終了する。

#### 【0053】

次に、I O 障害「無」と判定されステップ S607 へ移行すると、所定の瞬断監視時間 (図 6 に示す例では 1 秒前) の間に実行された I O 処理における、I O 障害の有無を判定する (S607)。その結果、過去 1 秒間に I O 障害が発生していないと判定されると、瞬断エラーに関する処理 (S608～S611) を実行することなく、エラーは発生していないと判定して (ステップ S612)、この処理を終了する。

#### 【0054】

一方、過去 1 秒間に I O 障害が発生していると判定されると、ステップ S603 へ移行して、瞬断エラーに関する処理 (S608～S611) を実行する。

#### 【0055】

瞬断エラーに関する処理では、まず、障害管理テーブル内の当該パスの瞬断エラー欄における開始時刻と監視期間を参照して、現在時が所定の監視時間範囲内であるかによって、所定の監視時間内に生じた I O 障害であるかを判定する (S608)。その結果、所定

の監視時間内に生じた I O 障害であれば、ステップ S 6 1 1 に移行し、瞬断エラーの検出値にエラー回数（1 回）を加算し、この処理を終了する。

【0056】

一方、ステップ S 6 0 8 による判定の結果、所定の監視時間外に発生したエラーであると判定されると、現在時刻を監視の開始時刻に設定し（S 6 0 9）、瞬断エラーの検出値を” 0 ”にして（S 6 1 0）、新たな監視時間を開始する。そして、瞬断エラーの検出値にエラー回数（1 回）を加算し（S 6 1 1）、この処理を終了する。

【0057】

図 7 は、本発明の第 1 の実施の形態のバス障害管理テーブルの説明図である。

【0058】

バス障害管理テーブルには、バス毎、エラーの種類毎に、当該種類のエラーに関する監視開始時刻、監視期間、判定値、検出値が記憶されている。

【0059】

例えば、瞬断エラーは、2003 年 9 月 1 日 10 時 20 分 30.02 秒から 30 日間を監視時間として監視が行われる。その監視時間内に瞬断エラーの発生が検出されると検出値が 1 ずつ加算され、障害管理テーブルが更新される。また、監視時間内に、判定値（20 回）の瞬断エラーの発生が検出されると、当該バスは瞬断エラーであると判定される。なお、監視時間内に複数回の瞬断エラーが検出されたときに瞬断エラーが発生したと判定するのは、電源の不安定等の一時的な要因によるバスの不安定を瞬断エラーであると誤認することを避けるためである。

【0060】

なお、性能エラーは複数回の I O 処理の平均によって定められるので、1 回でも性能の低下が認められると性能エラーであると判定することから、監視期間は定められていない。

【0061】

図 8 は、本発明の第 1 の実施の形態の稼動統計テーブル更新処理のフローチャートであり、サーバ切替処理（図 5）のステップ S 5 0 4 において実行される。

【0062】

まず、サーバ内に設けられたタイマから月日及び時間を取得して、当該月の特定の日時（図 6 に示す例では月初め、すなわち 1 日の 0 時）であるかを判定する（S 8 0 1）。そして、月初めでなければステップ S 8 0 4 に移行する。一方、月初めが検出されると、稼動統計テーブルに新たに始まる月のデータを記録する欄を追加し（S 8 0 2）、I O 処理において当該バスに送付したデータ量、及び当該 I O 処理に要した時間を” 0 ”に初期設定して（S 8 0 3）、新たな期間における性能の監視の準備をする。

【0063】

そして、正常に終了した I O 処理において使用したバスに送付したデータ量、及び当該 I O 処理に要した時間を稼動統計テーブルの該当する欄に加算する（S 8 0 4）。

【0064】

そして、所定の割合（図 6 に示す例では 60 %）だけ性能が劣化しているかを判定する（S 8 0 5）。この判定は、定期的（例えば、1 日 1 回）や、ユーザからの指示によって行われる。この判定の結果、60 %の性能劣化が見られたら性能エラーと判定して（S 8 0 6）、この処理を終了する。一方、60 %までも性能が劣化していなかったら性能エラーと判定することなく、この処理を終了する。

【0065】

なお、図 8 では、月単位で当該バスの性能（当該バスによる I O 処理のレスポンス）を集計しているが、任意の期間を設定して性能を監視してもよい。

【0066】

図 9 は、本発明の第 1 の実施の形態の稼動統計管理テーブルの説明図である。

【0067】

稼動統計管理テーブルには、バス毎、月毎に、I O 処理によって当該バスに送付したデ

ータ量の累積値、及び当該パスにおける I O 処理に要した時間の累積値が記録されている。そして、累積データ量を累積処理時間で除することによって、当該監視期間の I O 処理のレスポンスを求めることができる。

【0068】

なお、稼動統計テーブルには、性能が最も良好であった期間（通常は設置時）のレスポンスと、現在の監視期間の I O 処理のデータ量の累積値及び I O 処理の所要時間の累積値を記録していれば足りる。

【0069】

また、図 9 に示す稼動統計管理テーブルでは、I O 処理のデータ量と I O 処理の所要時間とを月毎に記録しているが、他の監視期間を定めて、パスの性能を監視してもよい。

【0070】

図 10 は、本発明の第 1 の実施の形態のサーバ切替判定処理のフローチャートであり、サーバ切替処理（図 5）のステップ S 506 において実行される。

【0071】

まず、障害管理テーブル（図 7）を参照して、パス毎に検出値が判定値を超えているかを判定する（S 1001）。検出値が判定値を超えていれば、エラーが発生したパスを特定する（S 1002）。この判定値との比較作業を全てのパスについて行う。

【0072】

そして、I O 処理負荷が高い時間帯（日中）であるかを判定する（S 1003）。本実施の形態のような情報処理システムでは、I O 処理負荷が高い時間帯（高負荷時間帯）と、I O 処理負荷が低い時間帯（低負荷時間帯）とが存在する。例えば、銀行のオンラインシステムや航空会社における座席予約システムの場合には日中の負荷が高くなるので、予め I O 処理負荷を調査した結果に基づいて、高負荷時間帯と低負荷時間帯とを定めておく。

【0073】

そして、日中（高負荷時間帯）であれば、サーバ切替閾値テーブル（図 11）の日中欄を選択し（S 1004）、夜間（低負荷時間帯）であれば、サーバ切替閾値テーブルの夜間欄を選択して（S 1005）、判定閾値を定める。図 11 に示すサーバ切替閾値テーブルでは、エラーの種類毎に全パス数に対して判定閾値となるパス数の比率が定められている。

【0074】

そして、ステップ S 1002 で特定したエラーが発生しているパス数と、S 1004 又は S 1005 で選択した閾値とを比較して（S 1006）、エラーパス数が閾値以上であれば、サーバ切替指示を発行する（S 1007）。

【0075】

すなわち、図 11 に示すサーバ切替閾値テーブルで、全パス数が 8 本であり、瞬断エラーの場合には、日中は 4 本（ $8 \text{本} \times 0.5 = 4$ ）が閾値となるが、夜間は 3 本（ $8 \text{本} \times 0.4 = 3.2$ ）が閾値となる。よって、夜間は日中より少ない本数のパスで瞬断エラーが検出されてもサーバの切替（フェイルオーバー）が発生することになる。よって、I O 負荷が高い時間帯（日中）にはフェイルオーバーを抑制して、I O 負荷が低い時間帯（夜間）に優先してフェイルオーバーをさせることができる。

【0076】

なお、判定閾値を時間によって定めるのではなく、稼動統計管理部 135 が収集した I O 処理量やスループットに基づいて、I O 処理負荷が低い時間帯の判定閾値を緩和して定め、フェイルオーバーを促進し、I O 処理負荷が高い時間帯の判定基準を厳しく定め、フェイルオーバーを抑制するようにしてもよい。

【0077】

図 12 は、本発明の第 2 の実施の形態のパスの障害判定に用いるテーブルの例を説明する図である。

【0078】

第2の実施の形態では、エラーが発生しているパスの、エラーの種類によって定められた点数を加算して、各パスで合計することによって点数を算出して、フェイルオーバをするかを判定するものであり、第1の実施の形態(図4)のようにパスに1種類のエラーが発生している場合の他、パスに複数種類のエラーが発生している場合も的確に障害を判定することができる。

#### 【0079】

具体的には、瞬断エラーの場合は”2”を、累計エラーの場合は”1.43”を、性能エラーの場合は”1.25”を加算して、合計点数を求める。そして、全パス本数と比較してフェイルオーバをするかを判定する。このエラーの種類毎のエラー点数は、スループットの低下に対する影響の大きさによって、そのエラーが重大であれば点数を高くして、早めにサーバを切り替えるようにしている。

#### 【0080】

例えば、テーブルの上段に示すケース4では、3本のパスに瞬断エラーが発生しており、2本のパスに累計エラーが発生しているので、合計点数は、 $3 \times 2 + 2 \times 1.43 = 8.86$  となり、全パス数(8本)を超えるのでフェイルオーバが必要と判定される。

#### 【0081】

同様に、テーブルの中段に示すケース5では、3本のパスに累計エラーが発生しており、4本のパスに性能エラーが発生しているので、合計点数は、 $3 \times 1.43 + 4 \times 1.25 = 9.29$  となり、全パス数(8本)を超えるのでフェイルオーバが必要と判定される。

#### 【0082】

さらに、テーブルの下段に示すケース6では、2本のパスに瞬断エラーが発生しており、2本のパスに累計エラーが発生しており、2本のパスに性能エラーが発生しているので、合計点数は、 $2 \times 2 + 2 \times 1.43 + 2 \times 1.25 = 9.36$  となり、全パス数(8本)を超えるのでフェイルオーバが必要と判定される。

#### 【0083】

図13は、本発明の第2の実施の形態のサーバ切替判定処理のフローチャートであり、サーバ切替処理(図5)のステップS506において実行される。なお、第2の実施の形態では、サーバ切替処理が異なるのみで他の処理は前述した第1の実施の形態と同じなので、その詳細な説明は省略する。

#### 【0084】

まず、障害管理テーブル(図7)を参照して、パス毎に検出値が判定値を超えているか否かについて判定する(S1301)。検出値が判定値を超えていれば、閾値を超えているエラーの種類を求め、エラーの種類を特定する(S1302)。この判定値との比較作業を全てのパスについて行う。

#### 【0085】

このとき、複数の種類のエラーの検出値が判定値を超えていれば、その中で一番重要なエラーをそのパスのエラーの種類として特定する。エラーの重要度は、スループットの低下に対する影響等を考慮して予め定めておく。例えば、瞬断エラー、累計エラー、性能エラーの順に重要なエラーとして定めておくといよい。また、検出された複数の種類のエラーのポイントを加算してもよい。

#### 【0086】

そして、IO処理負荷が高い時間帯(日中)であるか否かを判定する(S1303)。本実施の形態のような情報処理システムでは時間帯によってIO処理負荷が異なることから、高負荷時間帯と低負荷時間帯とを定めておく。

#### 【0087】

そして、日中(高負荷時間帯)であれば、エラー点数テーブル(図14)の日中欄を選択し(S1304)、夜間(低負荷時間帯)であれば、エラー点数テーブルの夜間欄を選

択して (S1305)、各エラーに関連して加算される点数を定める。図14に示すエラー点数テーブルでは、エラーの種類毎に、そのエラーが発生したパスに付与される点数が定められている。

#### 【0088】

そして、ステップS1302でパス毎に特定されたエラーの種類に対応して定まるエラー点数の合計値と、閾値（本実施の形態ではパス数）とを比較して (S1306)、エラー点数の合計値が閾値以上であれば、サーバ切替指示を発行する (S1307)。

#### 【0089】

以上説明したように、第2の実施の形態では、エラーの種類によって定められた点数を用いることによって、パスに発生したエラーが複数種類にわたっている場合にも的確に障害を判定することができる。また、8本のパス中の6本のパスに性能エラーが発生した場合、日中ではエラー点数が7.5 (6本×1.25) となり、夜間ではエラー点数が8.28 (6本×1.38) となり判定閾値 (8本) を超える。よって、I/O負荷が高い時間帯 (日中) にはフェイルオーバを抑制して、I/O負荷が低い時間帯 (夜間) に優先してフェイルオーバをさせることができる。

#### 【0090】

なお、エラー点数を時間によって定めるのではなく、稼動統計管理部135が収集したI/O処理量やスループットに基づいて、I/O処理負荷が低い時間帯はエラー点数を高くし、I/O処理負荷が高い時間帯はエラー点数を低く定めるようにしてもよい。

#### 【0091】

図15は、本発明の第3の実施の形態のサーバ切替判定処理のフローチャートであり、サーバ切替処理 (図5) のステップS506において実行される。

#### 【0092】

第3の実施の形態では、時間帯 (又は、I/O処理負荷) によって判定閾値を変える第1の実施の形態と異なり、時間帯によって判定対象となるエラー種別を変えることによって、フェイルオーバを抑制する。なお、第3の実施の形態では、サーバ切替処理が異なるのみで他の処理は前述した第1の実施の形態と同じなので、その詳細な説明は省略する。

#### 【0093】

まず、障害管理テーブル (図7) を参照して、パス毎に検出値が判定値を超えているか否かについて判定する (S1501)。検出値が判定値を超えていれば、閾値を超えているエラーの種類を求め、エラーの種類を特定する (S1502)。この判定値との比較作業を全てのパスについて行う。このとき、一つのパスにおいて複数の種類のエラーの検出値が判定値を超えていれば、当該複数のエラーをそのパスのエラーの種類として特定する。

#### 【0094】

そして、I/O処理負荷が高い時間帯 (日中) であるか否かを判定する (S1503)。本実施の形態のような情報処理システムでは時間帯によってI/O処理負荷が異なることから、高負荷時間帯と低負荷時間帯とを定めておく。

#### 【0095】

そして、日中 (高負荷時間帯) であれば、性能エラーを除外して、パス毎のエラー種別を特定し直す (S1504)。すなわち、あるパスに性能エラーのみが発生している場合には、当該パスにはエラーが発生していないものとする。また、あるパスに性能エラー及び他のエラー (例えば、瞬断エラー) のみが発生している場合には、当該パスには当該他のエラー (瞬断エラー) が発生しているものとする。

#### 【0096】

そして、ステップS1002で特定したエラーが発生しているパス数と、予め定めた閾値 (例えば、パス数の半分) とを比較して (S1505)、エラーパス数が閾値以上であれば、サーバ切替指示を発行する (S1506)。

#### 【0097】

なお、ステップS1502、S1504でパス毎に特定されたエラーの種類に対応して



定まるエラー点数の合計値と、閾値（本実施の形態ではパス数）とを比較して、エラー点数の合計値が閾値以上であれば、サーバ切替指示を発行してもよい。

#### 【0098】

以上説明したように、第3の実施の形態では、一部の種類のエラーについてはエラーの判定をしないので（又は、一部の種類のエラーのみについてエラーの判定をするので）、I/O負荷が高い時間帯（日中）にはフェイルオーバを抑制して、I/O負荷が低い時間帯（夜間）に優先してフェイルオーバをさせることができる。

#### 【0099】

本発明の実施の形態では、情報処理装置は、データ入出力要求の結果からパスに発生した障害を検出する障害検出部と、所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも、前記記憶装置に接続される情報処理装置を切り替える情報処理装置の切り替えを実行するかを判定する切替評価部と、前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求する情報処理装置を切り替える切替部と、を備えるので、サーバを切り替える際のパス障害検出期間中のスループットを向上させることができると共に、全パスに障害が検出される前に予防的にサーバを切り替えることによって、検出期間中のスループットの低下を防止することができる。

#### 【0100】

また、前記サーバ切替評価部は、前記障害検出部によって特定された障害の種類によって、前記情報処理装置の切り替えの実行の判定基準となるパス数を変更するので、障害の重大度に応じて適切な情報処理装置の切り替えの判定をすることができる。

#### 【0101】

また、前記サーバ切替評価部は、前記障害検出部によって特定された複数種類の障害の組み合わせに基づいて前記情報処理装置の切り替えの実行を判定するので、複数のパスに異なる種類の障害が発生しても適切に情報処理装置の切り替えの判定をすることができる。

#### 【0102】

また、前記サーバ切替評価部は、データ入出力要求の負荷状態に応じて前記情報処理装置の切り替えの実行の判定基準となるパス数を変更するので、高負荷時間帯における情報処理装置の切り替えを抑制し、低負荷時間帯における情報処理装置の切り替えを促進するので、業務に影響が少ない時間帯に情報処理装置の切り替えをさせることができる。

#### 【図面の簡単な説明】

#### 【0103】

【図1】本発明の第1の実施の形態の情報処理システムの全体構成を示すブロック図である。

【図2】本発明の第1の実施の形態のサーバ100の詳細を表した機能ブロック図である。

【図3】本発明の第1の実施の形態の障害検出部133の動作を示すブロック図である。

【図4】本発明の第1の実施の形態のクラスタ切替判定に用いるテーブルの説明図である。

【図5】本発明の第1の実施の形態のサーバ切替処理のフローチャートである。

【図6】本発明の第1の実施の形態の障害管理テーブル更新処理のフローチャートである。

【図7】本発明の第1の実施の形態の障害管理テーブルの説明図である。

【図8】本発明の第1の実施の形態の稼動統計テーブル更新処理のフローチャートである。

【図9】本発明の第1の実施の形態の稼動統計管理テーブルの説明図である。

【図10】本発明の第1の実施の形態のサーバ切替判定処理のフローチャートである。

【図11】本発明の第1の実施の形態のサーバ切替閾値テーブルの説明図である。

【図 1 2】 本発明の第 2 の実施の形態のパスの障害判定に用いるテーブルの例の説明図である。

【図 1 3】 本発明の第 2 の実施の形態のサーバ切替判定処理のフローチャートである。

【図 1 4】 本発明の第 2 の実施の形態のエラー点数テーブルの説明図である。

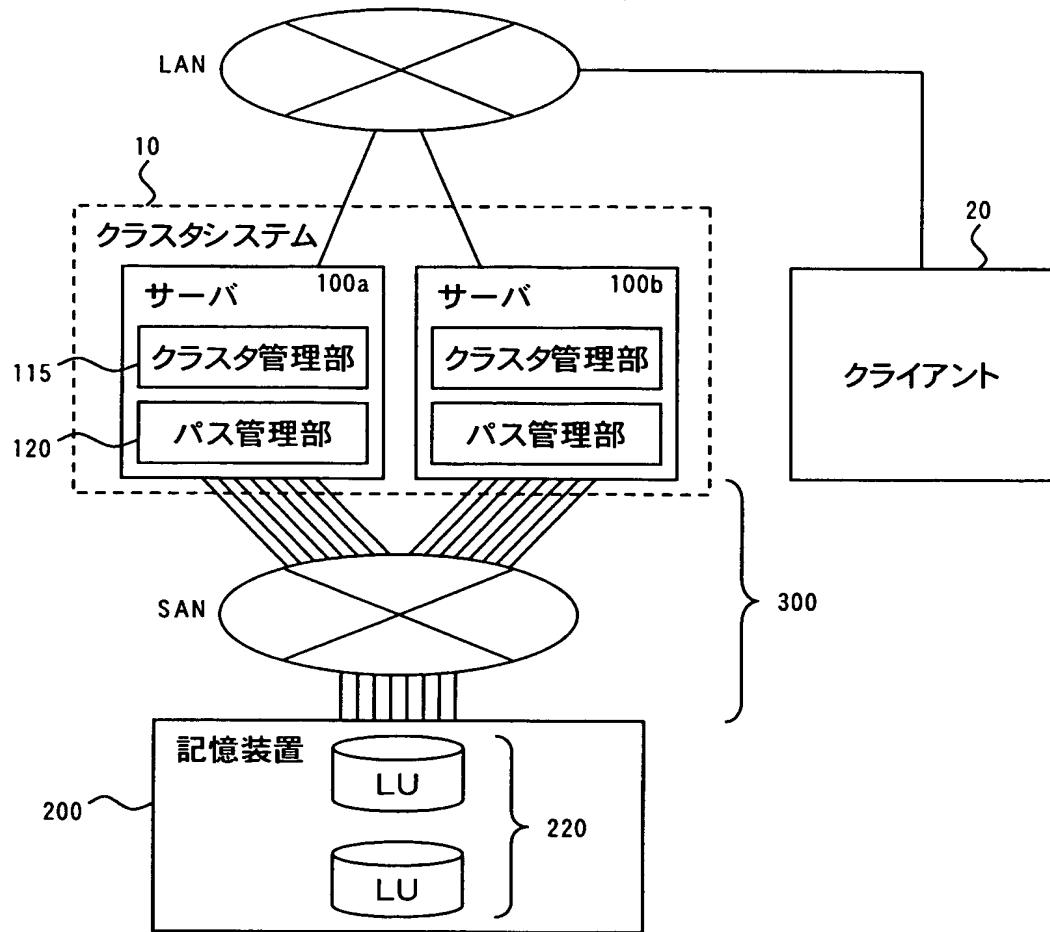
【図 1 5】 本発明の第 3 の実施の形態のサーバ切替判定処理のフローチャートである。

【符号の説明】

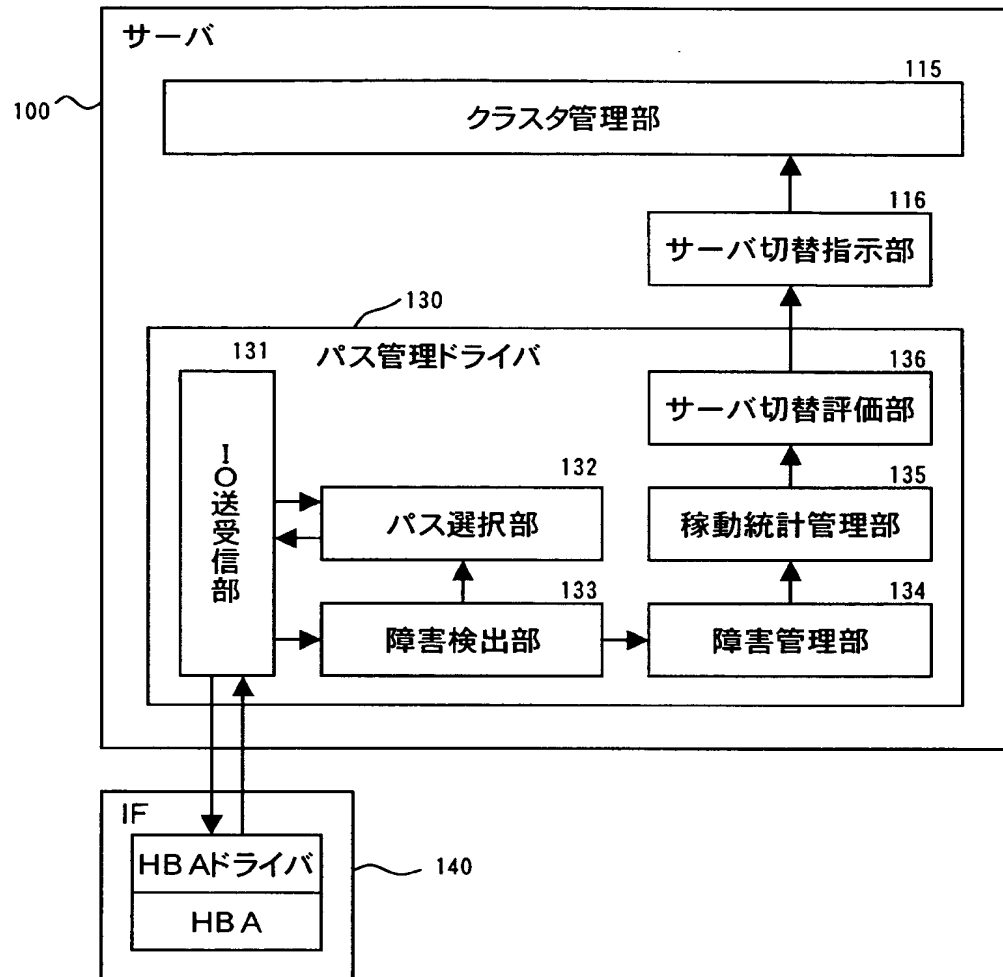
【 0 1 0 4 】

1 0	クラスタシステム
1 0 0	情報処理装置
1 1 0	アプリケーション
1 1 5	クラスタ管理部
1 1 6	サーバ切替部
1 2 0	パス管理部
1 3 0	パス管理ドライバ
1 3 1	I O 送受信部
1 3 2	パス選択部
1 3 3	障害検出部
1 3 4	障害管理部
1 3 5	稼動統計管理部
1 3 6	サーバ切替評価部
1 4 0	インターフェース ( I F )
2 0 0	記憶装置
2 1 0	ディスク制御部
2 2 0	論理ユニット ( L U )
3 0 0	物理パス

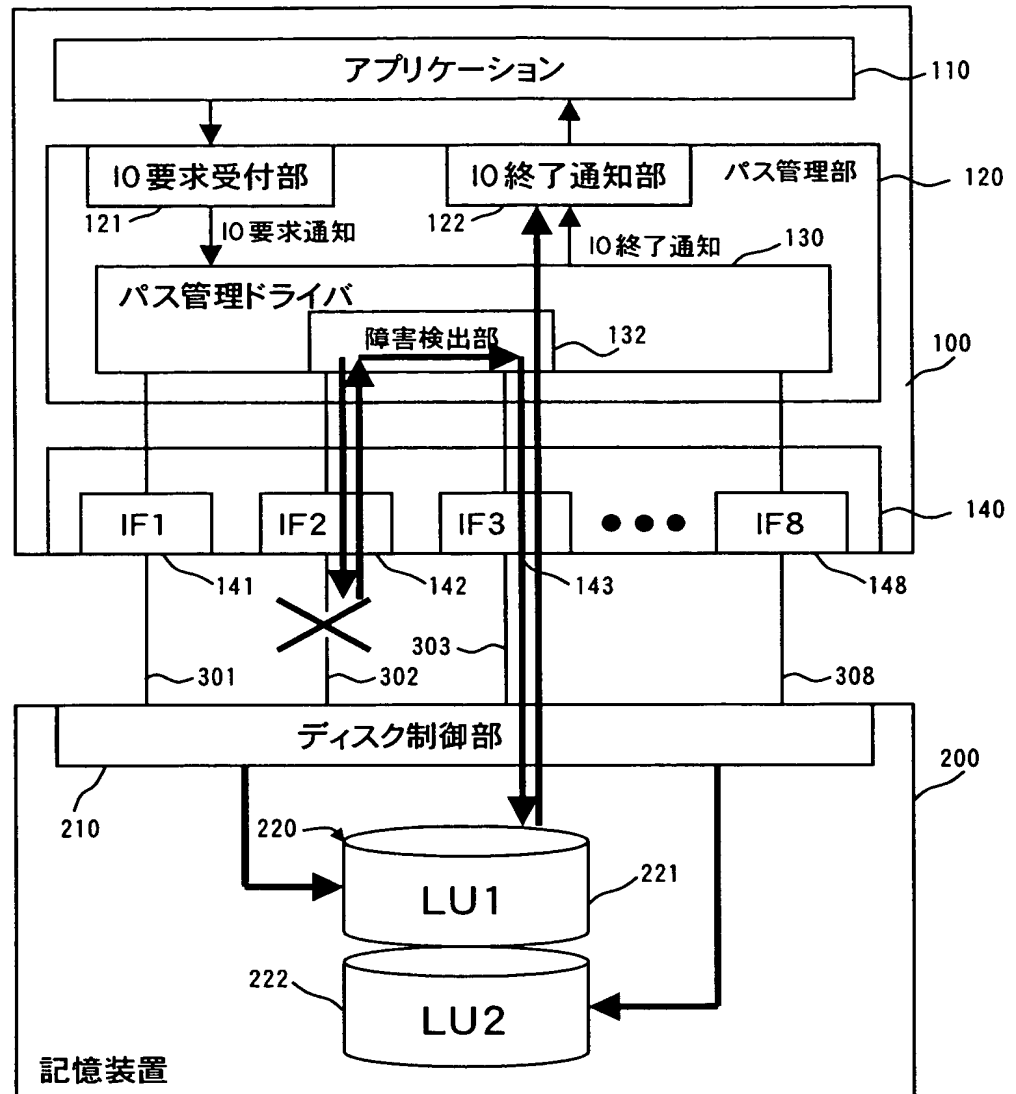
【書類名】 図面  
【図 1】



【図 2】



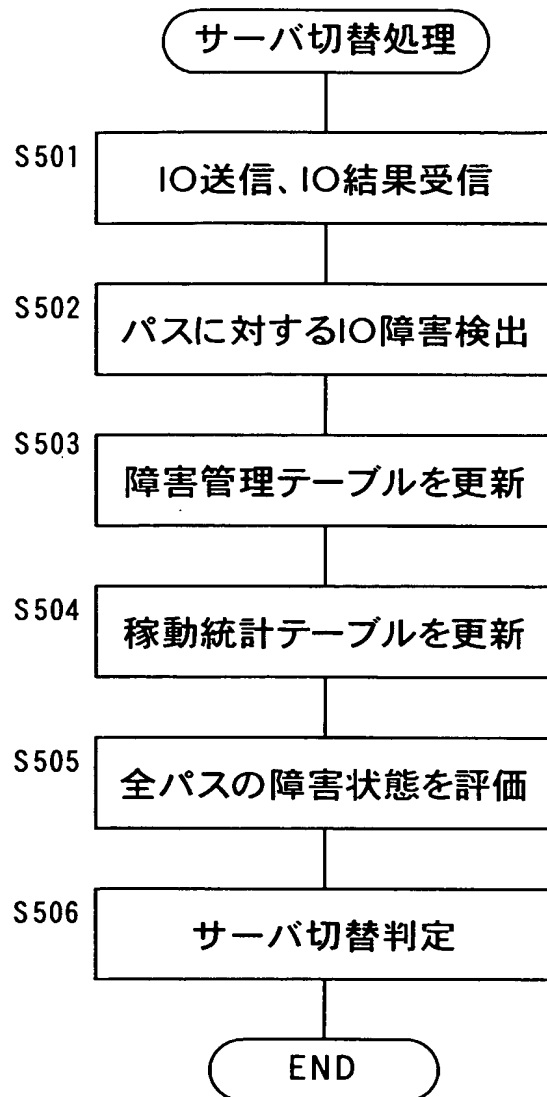
【図3】



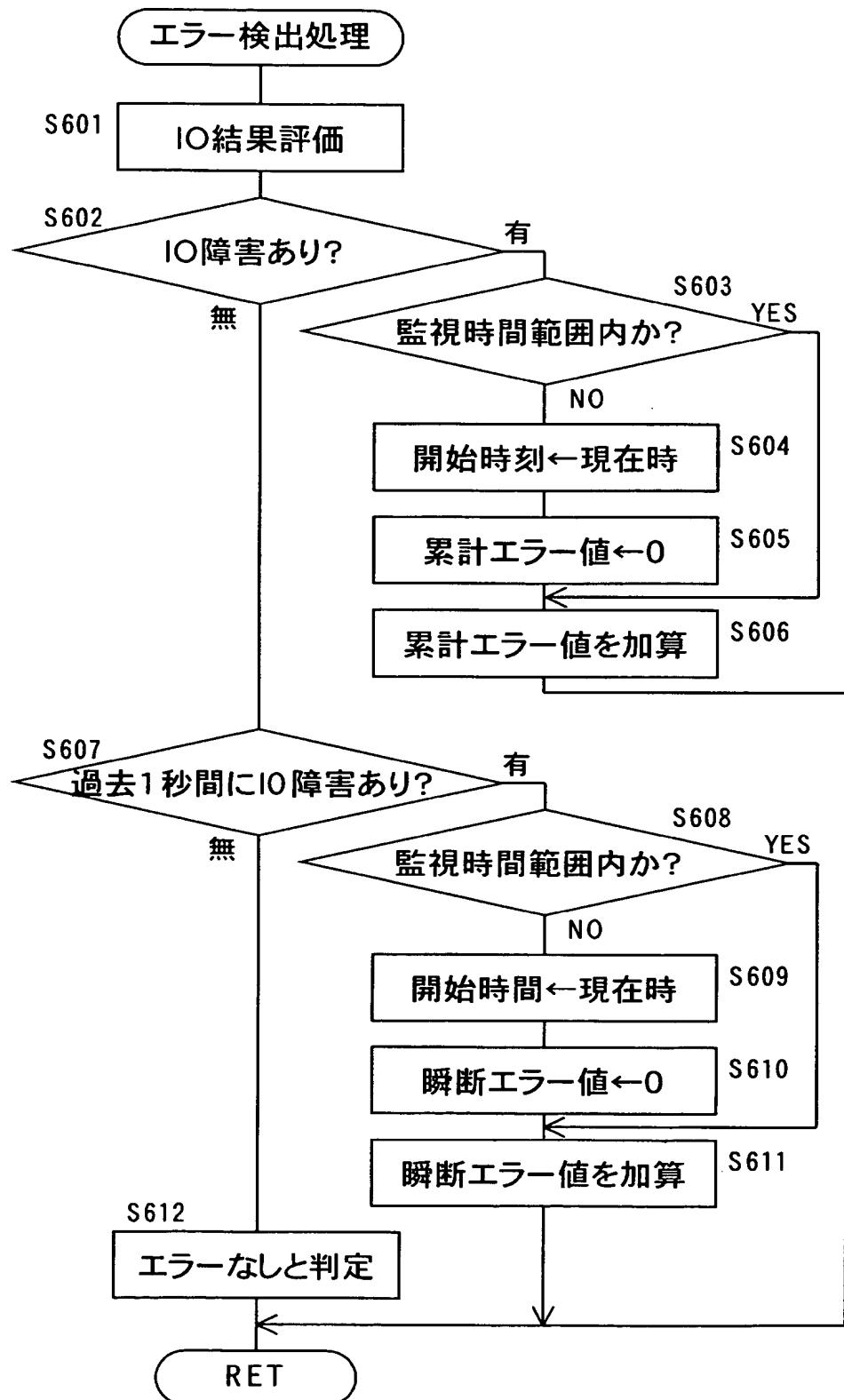
【図 4】

ケース	パス1	パス2	パス3	パス4	パス5	パス6	パス7	パス8
1	瞬断	瞬断	瞬断	瞬断	正常	正常	正常	正常
2	累計	累計	累計	累計	累計	累計	正常	正常
3	性能	性能	性能	性能	性能	性能	性能	正常

【図 5】



【図6】



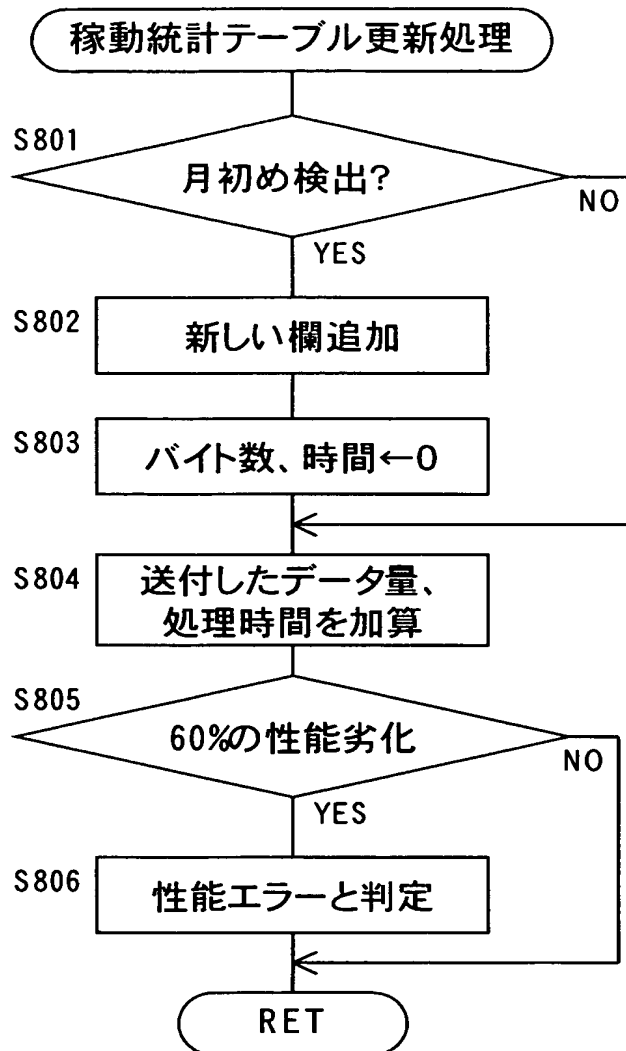


【図 7】

パス	エラーの種類	開始時刻	監視期間	判定値	検出値
1	瞬断エラー	2003/9/1 10:20:30.020	30	20	19
	累計エラー	2003/9/1 10:20:30.020	30	50	30
	性能エラー	2003/9/1 10:20:30.020	-	1	0
2	瞬断エラー	2003/9/1 10:20:30.020	30	20	0
	累計エラー	2003/9/1 10:20:30.020	30	50	0
	性能エラー	2003/9/1 10:20:30.020	-	1	0
3	● ● ●				

バス障害管理テーブル

【図 8】

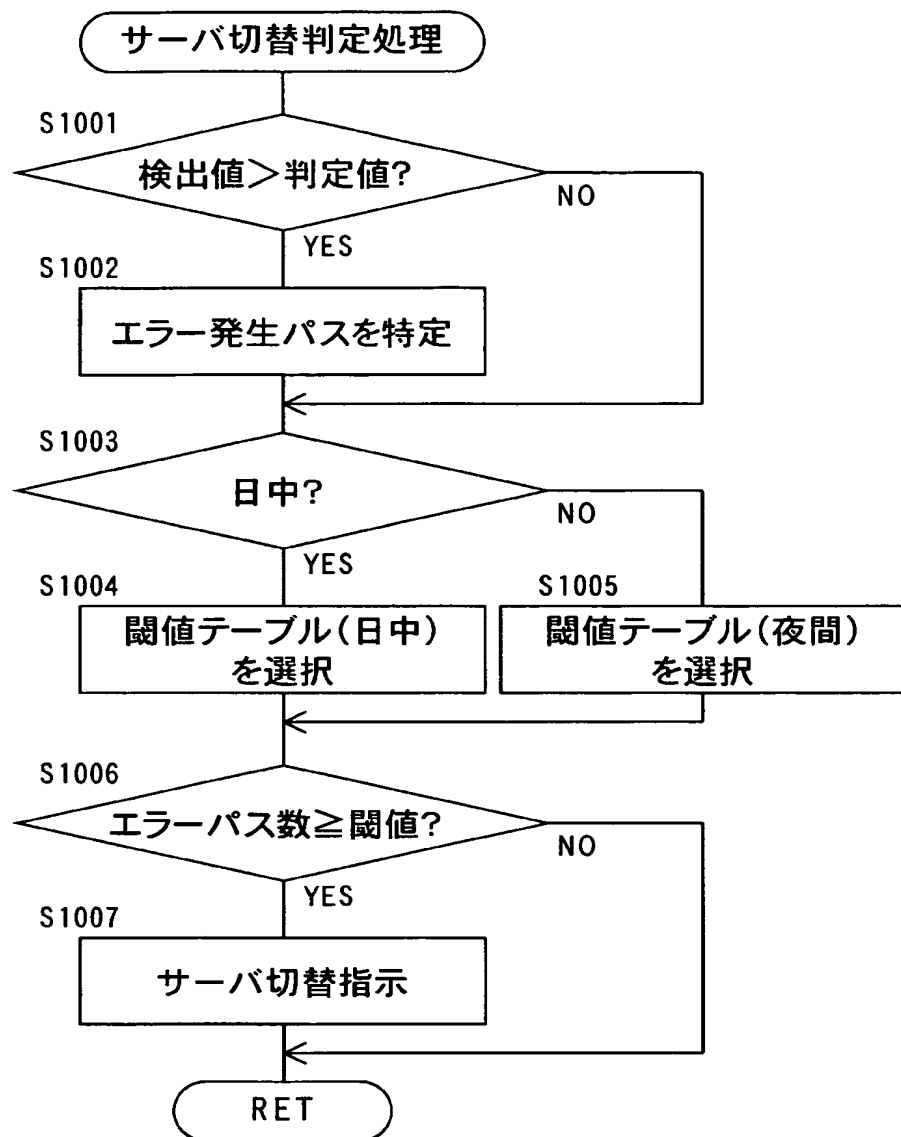


【図 9】

パス	2003/8/1		2003/9/1		2003/10/1		2003/11/1	
	バイト数	時間	バイト数	時間	バイト数	時間	バイト数	時間
1	100	1000	90	1000	80	1000	50	1000
2	100	1000	90	1000	80	1000	50	1000
3	100	1000	90	1000	80	1000	50	1000
4	100	1000	90	1000	80	1000	50	1000
5	100	1000	90	1000	80	1000	50	1000
6	100	1000	90	1000	80	1000	50	1000
7	100	1000	90	1000	90	1000	90	1000
8	100	1000	90	1000	90	1000	90	1000

稼動統計管理テーブル

【図 10】



【図 1 1】

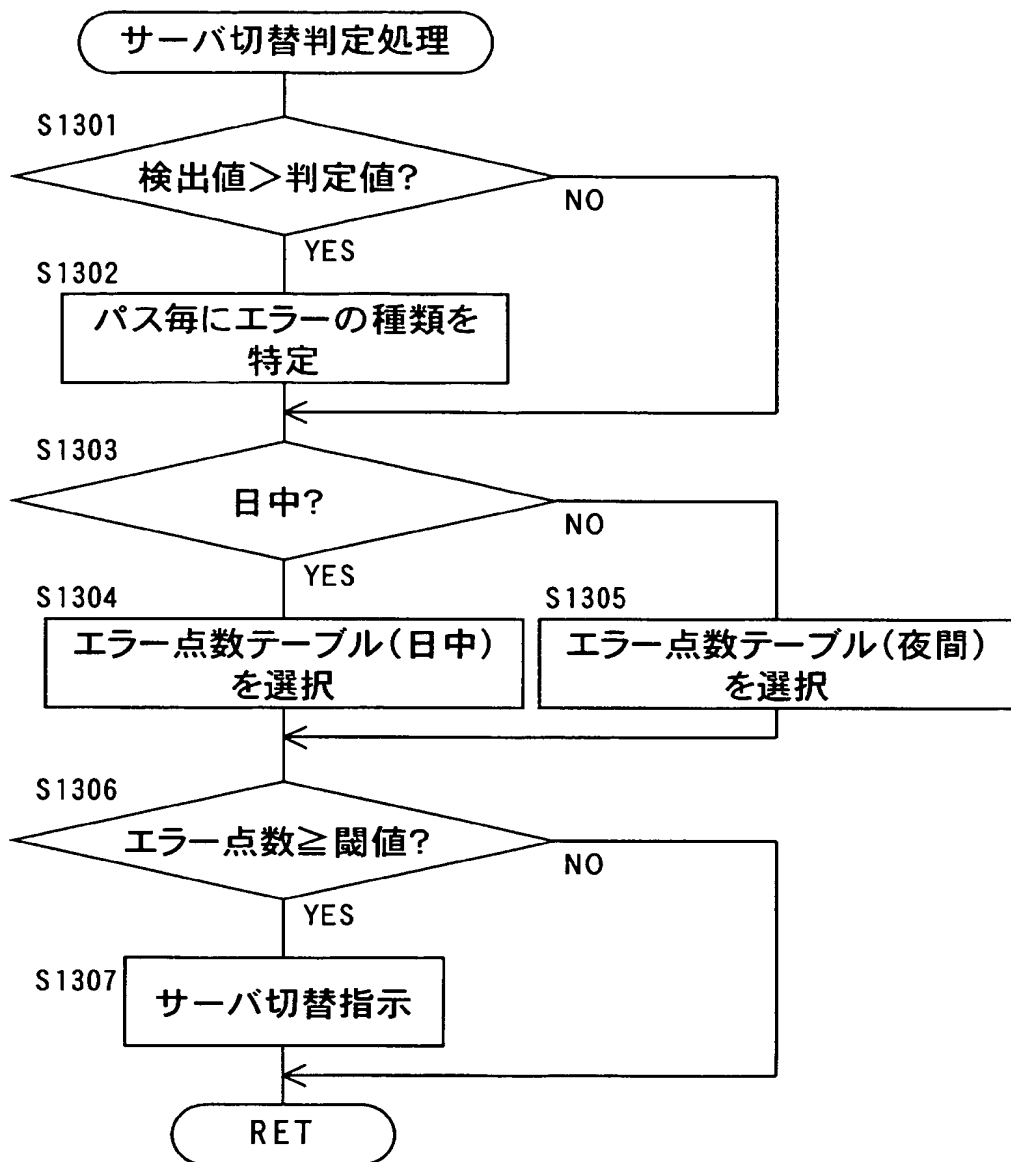
エラーの種類	パス数比率	
	日中	夜間
瞬断エラー	50%	40%
累計エラー	80%	70%
性能エラー	90%	80%

サーバ切替閾値テーブル

【図 1 2】

ケース	パス1	パス2	パス3	パス4	パス5	パス6	パス7	パス8
4	瞬断	瞬断	瞬断	累計	累計	正常	正常	正常
5	累計	累計	累計	性能	性能	性能	性能	正常
6	瞬断	瞬断	累計	累計	性能	性能	正常	正常

【図 13】



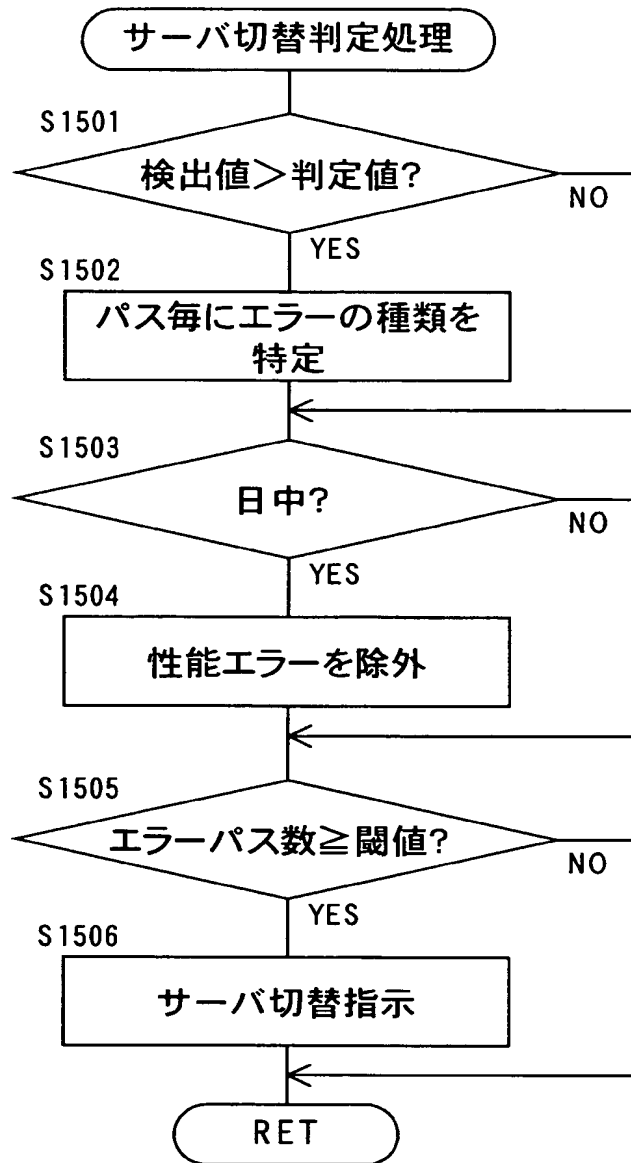
【図 1 4】

エラーの種類	点数	
	日中 (1)	夜間 (1. 1)
瞬断エラー	2	2. 2
累計エラー	1. 43	1. 57
性能エラー	1. 25	1. 38

エラー点数テーブル



【図 15】



【書類名】 要約書

【要約】

【課題】 全パスに障害が検出される前に、予防的にサーバを切り替えることによって検出期間中のスループットの低下を防止する。

【解決手段】 記憶装置と、前記記憶装置に対してデータ入出力を要求する複数の情報処理装置と、を備える情報処理システムにおいて、前記情報処理装置は、データ入出力要求の結果からパスに発生した障害を検出する障害検出部と、所定数のパスに障害が発生したことを検出すると、全てのパスに障害が発生する前でも、前記記憶装置に接続される情報処理装置を切り替えるフェイルオーバを実行するかを判定する切替評価部と、前記切替評価部の判定結果に基づいて、前記論理ユニットに対してデータ入出力を要求する情報処理装置を切り替える切替部と、を備える。

【選択図】 図 2

特願 2 0 0 3 - 3 8 7 9 4 2

出 願 人 履 歴 情 報

識別番号 [ 0 0 0 0 0 5 1 0 8 ]

1. 変更年月日 1 9 9 0 年 8 月 3 1 日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台 4 丁目 6 番地

氏 名 株式会社日立製作所